

Very Deep VAEs Generalize Autoregressive Models and Can Outperform Them on Images

Year: 2020 | Citations: 380 | Authors: R. Child

Abstract

We present a hierarchical VAE that, for the first time, outperforms the PixelCNN in log-likelihood on all natural image benchmarks. We begin by observing that VAEs can actually implement autoregressive models, and other, more efficient generative models, if made sufficiently deep. Despite this, autoregressive models have traditionally outperformed VAEs. We test if insufficient depth explains the performance gap by scaling a VAE to greater stochastic depth than previously explored and evaluating it on CIFAR-10, ImageNet, and FFHQ. We find that, in comparison to the PixelCNN, these very deep VAEs achieve higher likelihoods, use fewer parameters, generate samples thousands of times faster, and are more easily applied to high-resolution images. We visualize the generative process and show the VAEs learn efficient hierarchical visual representations. We release our source code and models at <https://github.com/openai/vdvae>.